医保等行政管理数据确定慢性病洗脱期最佳策略的系统综述

杨文怡，王敬鑫，艾丽梅，等. 医保等行政管理数据确定慢性病洗脱期最佳策略的系统综述［J］. 中国全科医学，2023. ［Epub ahead of print］. DOI：10.12114/j.issn.1007-9572.2023.0005

杨文怡　王敬鑫 艾丽梅　万霞*

1. 100005 北京市，中国医学科学院基础医学研究所 北京协和医学院基础学院 流行病与卫生统计学系

*通讯作者：万霞，博士生导师，研究员；wanxiasnake@163.com

【摘要】 当使用医疗保险等行政管理数据时，为正确确定反复就医的慢性病新发病时点，研究者尝试使用洗脱期为基础确定新发病例。本文系统地检索了在 PubMed、Web of science、Embase、中国知网、维普、万方数据库中从建库至 2022 年 10 月 1 日发表的所有利用行政管理数据探究慢性疾病发病、患病的相关文献，筛选并提取相关信息，总结洗脱期时长的确定方法。 截止 2022 年 10 月 1 日，共纳入 26 篇有效文献，结果显示，文献来源主要集中在加拿大、美国、澳大利亚等行政管理数据完整丰富的国家，研究疾病包括糖尿病、肿瘤、精神分裂等多种慢性疾病。研究认为，设定合适的洗脱期是准确识别发病病例的基础。目前文献中确定洗脱期的方法主要分为三大类，包括经验确定法、一致性检验法和逆向生存函数法。三种方法均有相应的优势和局限性，方法的选择、判断标准和稳定性有待进一步的探究。

【关键词】 慢性疾病 行政管理数据 医疗保险数据 洗脱期 发病 患病

A systematic review of optimal strategies for identifying chronic disease washout period with administrative data like insurance data

YANG Wen-yi, WANG Jing-xin, AI Li-mei, ZHAO Jian , WAN Xia

1. Institute of Basic Medical Sciences, Chinese Academy of Medical Science & School of Basic Medicine, Peking Union Medical College, Beijing 100005, China

**Corresponding authors: Wan Xia, Ph.D.supervisor, Researcher; wanxiasnake@163.com*

【Abstract】 When using administrative data such as medical insurance, the researchers attempted to identify new cases based on the washout period in order to correctly identify the new onset point of chronic diseases with repeated visits. In this paper, all literatures on exploring the incidence and illness of chronic diseases using administrative data from PubMed, Web of science, Embase, CNKI, VIP and Wanfang databases published from the establishment of the database to October 1, 2022 were systematically searched. Relevant information was screened and extracted, and the determination method of washout duration was summarized. As of October 1, 2022, a total of 26 valid literatures were included. The results showed that the literatures were mainly from Canada, the United States, Australia and other countries with complete and abundant administrative data, and the research diseases included diabetes, cancer, schizophrenia and other chronic diseases. It is believed that setting an appropriate washout period is the basis for accurate identification of cases. At present, the methods for determining washout period in literature mainly fall into three categories, including empirical determination method, consistency test method and reverse survival function method. The three methods have corresponding advantages and limitations, and the selection, judgment criteria and stability of the methods need to be further explored.

【Key words】 Chronic disease; Administrative data; Insurance data; Washout period; Incidence; Prevelence

背景

为了有效预防和控制糖尿病、肿瘤、慢性阻塞性肺病（COPD）等多种慢性疾病，我们需要准确估计疾病的患病和发病率。目前有一些大型队列、调查研究对疾病的患病、发病展开精确调查，但这种研究方式耗时长、耗资大。因此，考虑应用医保等行政管理收集的数据进行慢性疾病流行病学特征分析，但这些数据往往只能获取有限年份的记录，不能提供病人从出生到死亡全部生命周期的医疗记录，且慢性病病人通常会反复就医，从这些截断的重复记录中有效识别出新发病例有一定难度。为解决这一问题，研究者们一般使用回顾期（look-back period），也叫洗脱期（washout period）为基础确定新发病例。其涵义指的是患者在目标检索年份内有某慢性病的登记记录，即为可疑的发病病例，以检索年为基础，向前回顾，在限定的洗脱期内没有该病记录即确定为新发病例。洗脱期的长短目前尚没有统一结论[1, 2]。洗脱期太短会造成发病率的高估，太长则会导致数据利用不[3]。一些研究者认为不同疾病有不同的发展轨迹和疾病特征[4-6]，在使用不同数据库确定不同疾病洗脱期时，均应采用合适的方法确定最佳洗脱期时长。本次研究系统综述了确定洗脱期时长的方法，以期为中国研究者后续使用行政管理数据限定慢性疾病洗脱期，正确识别发病病例提供思路。

1.　方法

1.1 数据来源及检索策略

研究检索了从建库至 2022 年 10 月 1 日发表的所有相关文献。研究数据来源于 3 个英文数据库（PubMed、

Web of Science、Embase）和 3 个中文数据库（CNKI、VIP、万方数据库）。中文检索词包括：医保数据、官方数据、行政数据、医院数据、患病，发病，洗脱期、回顾期、窗口期。英文检索词包括：administrative data、insurance data、hospital data、Medicare、prevalence、incidence，look-back period、wash-out time、clearance time、disease-free time、observation time。

1.2 纳排标准

纳入标准：（1）研究数据为行政管理数据，包括医疗保险数据、疾病登记注册数据、医院记录数据等；（2）研究疾病为慢性病，同一患者同种疾病出现多次记录；（3）研究目的为探究疾病患病、发病情况，聚焦流行病学特征；（4）研究中有洗脱期的相关内容。

排除标准：（1）非中英文数据（2）研究数据非官方数据，如队列调查、病例对照试验调查数据等；（3）研究疾病非慢性病，如感染性疾病、急性病等；（4）研究目的为评估药物有效性、治疗模式、费用负担等非疾病流行病学特征；（5）研究中未出现洗脱期相关；（6）非原始研究，如：评论、系统综述等；（7）会议摘要；（8）无法获取原文。

1.3 信息提取

由两名独立的研究人员对符合纳入标准的文献相关信息进行提取，并对提取信息进行交叉核对，以保证信息提取质量。提取信息包括：（1）文章基本信息：文章名称、第一作者、发表年份；（2）使用数据库基本信息：国家、数据库名称、涵盖人群范围、数据库截取年份；（3）研究设计：研究疾病、最佳洗脱期判断方法及标准、采用洗脱期时长。

1.4 研究方法

使用 EndnoteX20 建立数据库，严格依据纳入排除标准筛选文献。使用 Draw 软件绘制流程图，使用 Microsoft Excel 2010 对文章信息进行提取和基本特征分析，描述综述洗脱期确定方法。

2. 结果

2.1 文献筛检流程及结果

本次研究通过知网、万方、维普等数据库，检索出中文文献 21 篇；通过 PubMed、Embase、Web of Science 检索相关英文文献 1203 篇。经过初筛、复筛，筛选基于行政管理数据库评估慢性病患病或发病等流行病学特征的相关文献 54 篇，排除其中 28 篇未使用洗脱期研究，最终纳入 26 篇文献进入本次系统综述，文献具体筛检流程见图 1。

纳入的 26 篇文献均为英文文献，研究国家主要集中在加拿大、美国、澳大利亚等，研究疾病包括糖尿病、肿瘤、精神分裂等多种慢性疾病，研究目的聚焦于多种慢性疾病的发病、患病确定。所有纳入的研究中，多数研究（46.15%，n=12）文献数据来源为医疗保险数据，8 篇为医院病人注册记录，其余文献数据来源包括有居民健康项目注册、疾病调查项目注册等大型行政管理数据。 研究使用的数据库年份跨度较长，仅有两篇文献数据包含年份小于 5 年，半数以上（53.85%，n=14）的数据库年份跨度大于 10 年。纳入文献及使用数据库基本信息见表 1。

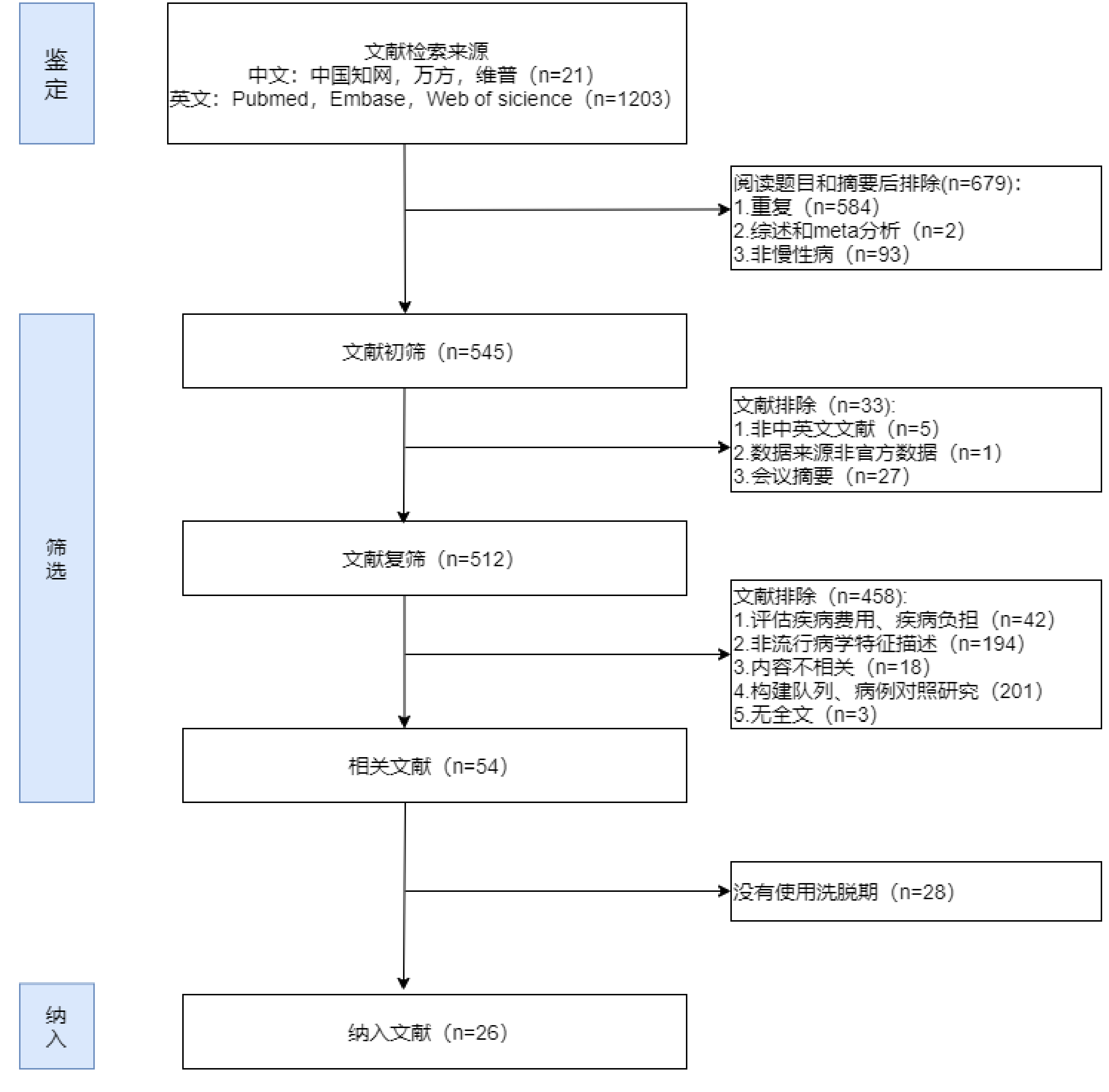

鉴定
文献检索来源
中文：中国知网，万方，维普（n=21）
英文：Pubmed，Embase，Web of sicience（n=1203）
阅读题目和摘要后排除(n=679)：
1.重复（n=584）
2.综述和meta分析（n=2）
3.非慢性病（n=93）
筛选
文献初筛（n=545）
文献排除（n=33）:
1.非中英文文献（n=5）
2.数据来源非官方数据（n=1）
3.会议摘要（n=27）
文献复筛（n=512）
文献排除（n=458）:
1.评估疾病费用、疾病负担（n=42）
2.非流行病学特征描述（n=194）
3.内容不相关（n=18）
4.构建队列、病例对照研究（201）
5.无全文（n=3）
相关文献（n=54）
没有使用洗脱期（n=28）
纳入
纳入文献（n=26）
图 1. 文献筛选流程图

表 1 文献及使用数据库基本信息

作者	发表年份	研究疾病	使用数据基本信息				
			国家/地区	名称	类型	使用年份	涵盖人群
K. M. Ødegaard	2022	心衰	挪威	The Norwegian Patient Registry	医院登记	2008-2018	挪威全体人群
K. J. Brameld	2003	糖尿病、心肌梗死	美国	Patient Registry	医院登记	1980-1995	—
N. Beaudet	2014	下背痛	加拿大魁北克	the Régie de l'Assurance Maladie du Québec (RAMQ)	医疗保险	2000-2007	魁北克 50%以上人群
S. Asghari	2009	糖尿病	加拿大魁北克	the Régie de l'Assurance Maladie du Québec(RAMQ)	医疗保险	1992-2002	5686255
Alain Vanasse	2012	精神分裂	加拿大魁北克	the Régie de l'Assurance Maladie duQue´bec (RAMQ).	医疗保险	1996-2006	魁北克 50%以上人群
E. I. Benchimol,	2014	慢性肠道疾病	加拿大安大略省	Hospital Discharge Abstract Data (DAD)、Canadian Institute for Health Information (CIHI-DAD)	医院登记	2002-2012	魁北克 50%以上人群
J.C.Lin	2014	糖尿病视网膜病变	中国台湾	Health Insurance Database of taiwan	医疗保险	2005-2011	台湾 99%人口
Gerhard Sulo	2015	心肌梗死	挪威	Norway project	其他	1994-2009	—
J. Czwikla	2017	乳腺癌、前列腺癌、结直肠癌	德国	Statutory Health Insurance	医疗保险	2006-2014	德国 90%人口
Rassen, J. A.	2019	慢阻肺、银屑病、精神分类、囊性纤维症	美国	Electronic Health Records	其他	2011-2014	美国大部分人群
Sascha Abbas	2012	糖尿病、结直肠癌、心衰	德国黑森州	Statutory Health Insurance sample (SHI)	医疗保险	2000-2008	
J. D. Stein	2013	青光眼、黄斑变性	美国明尼苏达	The Clinformatics Data Mart database	医疗保险	2001-2011	总共涵盖 2457 人

M. Rosenlund	2020	24 中慢性病	瑞典	National Patient Register	医院登记	2005-2013	7943807
N. Schmedt	2017	多发性硬化	德国	German Pharmaco epide miological Research Database (GePaRD)	其他	2004-2013	4175877
R. Camplain	2017	心衰	美国	the Atherosclerosis Risk in Communities (ARIC)	其他	2000-2012	15744
C. Ohlmeier	2019	亨廷顿氏病	德国	Applied Health Research Berlin (InGef) Research database	其他	2010-2016	约 6.7 百万人
B. K. Koo	2014	糖尿病、动脉粥样硬化	韩国	The National Health Insurance (NHI)	医疗保险	2007-2011	韩国 97%人群
M. Jürisson	2014	髋骨骨折	爱沙尼亚	Estonian Health Insurance Fund	医疗保险	2005-2012	爱沙尼亚 96%以上人口
A. S. Bukhbinder	2022	阿尔兹海默	美国	the Optum Clinformatics Data Mart (CDM)	其他	2009-2019	—
L. Zhou	2022	血栓形成	加拿大	Canadian Institute for Health Information's Discharge Abstract Database(CIHI-DAD)	医院登记	2011-2017	—
J. Katzenellenbogen	2019	风湿性心脏病	澳大利亚	ARF register data	医院登记	2000-2017	—
Ou.Y	2012	青光眼、阿尔兹海默症、其他痴呆	美国	Medicare 5% claims data	医疗保险	1994-2007	美国保险（Medicare）5%以上人口
B. B. Clissold	2017	中风	维多利亚	Victorian Admitted Episodes Dataset	医院登记	1997-2008	4,932,422
J. K. Ahn	2020	风湿性疾病	韩国	National Health Insurance Service	医疗保险	2007-2018	韩国 97%人群
K. K. Anderson	2012	精神分裂	加拿大魁北克	the Régie de l'Assurance Maladie du Québec(RAMQ)	医疗保险	2000-2006	魁北克 50%以上人群
M.Katzenellenbogen	2010	心肌梗死	澳大利亚	Hospital Morbidity Data Collection (HMDC)、Mortality Data Set	医院登记	1980-2004	—

2.2 最佳洗脱期探究方法

综述研究发现，目前文献中确定洗脱期的方法主要分为三大类，包括经验确定法、一致性检验法和逆向生存曲线法。纳入的研究中有 12 篇文献使用专家建议、临床经验、文献检索或受限于数据获取年份，直接确定洗脱期；9 篇文献使用一致性检验的相关指标，如 Kappa 值，阳性预测值（PPV），高估率等评估洗脱期对发病率估计的影响、确定洗脱期；3 篇文献使用逆向生存函数确定最佳洗脱期。

2.2.1 一致性比较法

纳入文献中有 9 篇文献采用一致性比较法，比较指标包括 Kappa 值、阳性预测值(Positive Predictive Values，PPVs)、过高估计率、阴性预测值、灵敏度、准确度。所有方法均采用数据可达到的最长洗脱期为“金标准”，使用以上指标定量比较不同洗脱期下确定的发病人数（Incident patient defined by different clearance period，IDD)，与“金标准”确定的发病人数(Incident patient defined by golden criterion，IDG)的差异性。其中有 6 篇文章采用过高估计率($\mathrm{overestimation} = (\mathrm{IDD} - \mathrm{IDG})/\mathrm{IDG}$)，2 篇文章采用 Kappa 值，3 篇文章采用 PPVs，1 篇文章采用灵敏度、特异度和阴性预测值为指标比较一致性。其中仅有 Kappa 值提出了相关确定判断标准，按照 Byrt、Shabnan[7]等人提出的分类方法，当 kappa 值介于-1-0 时，认为“没有一致性”；介于 0-0.20 时，认为“一致性不佳”；介于 0.21-0.40 时，认为“低一致性”；介于 0.41-0.6 时，认为“一般一致”；介于 0.61-0.80 时，认为“一致性较好”；介于 0.81-0.90 时，认为“一致性好”；介于 0.91-1.0 时，认为“一致性极好”。其他指标均没有确定标准判断一致性。

2.2.2 逆向生存函数法

逆向生存函数法借用生存函数的思路，一般使用 Kaplan-Meier 或寿命表法进行估计。在该方法中，终点时间为在洗脱期内出现重复记录，生存被定义为在给定的洗脱期内没有捕捉到该患者的重复记录，即可被定义为新发病例。具体来说，如计算 j 年中某慢性病 i 的发病人数，a 代表某病人在 j 年中第一次因疾病 i 被记录的日期，b 代表该病人洗脱期内离 a 最近的一次疾病 i 记录日期，c 代表洗脱期的开始日期。如果该病人在给定的洗脱期内有重复记录，该病人的”生存时间”为$l_x = a - b$；如果该病人在给定的洗脱期内没有重复记录，该病人则被定义为“删失”，即“生存”，删失时间为$l_c = a - c$，据此建立生存函数，计算在给定洗脱期 T 下，研究对象的生存概率[$S(t)$],一般可以使用 Kaplan-Meier 或寿命表法进行估计,并绘制生存函数曲线。同时，Kate J Brameld[8]等和 N. Beaudet[9]等在文章中也计算了风险函数[$\mathrm{h}(\mathrm{t})$]，其含义为生存时间已达到t的个体在接下来的Δt时间内发生终点事件的瞬时风险，并绘制风险函数曲线。研究认为，存在某一时点t_f使得$\mathrm{h}(t_f)$趋近于零，$S(t_f)$趋近于定值，该时点，即为最佳洗脱期。在实际研究中，Kate J Brameld[8]等认为当$\mathrm{h}(t_f) < 0.00001$ 时，即为趋近于 0，而 N. Beaudet[9]等则未给出具体标准。S. Asghari[10]等人的研究则未计算风险函数，仅在生存函数不再变动，趋近于稳定时，取值最佳洗脱期。

2.3 洗脱期时长

纳入的 26 篇文献，所研究的疾病包括糖尿病、中风、精神分裂等多种慢性疾病，其中心血管疾病研究文献最多（n=10），其次为糖尿病（n=4）。研究使用的洗脱期时长从 0.5—15 年不等，不同疾病洗脱期长短使用差异明显。同时，研究发现同一疾病的洗脱期长短差异也较大，4 篇研究糖尿病的文献中，使用的洗脱期时长最长为 13 年，最短为 1 年。纳入文献洗脱期确定方法及洗脱期时长详细信息见表 2。

表 2.洗脱期确定方法及时长

作者	发表年份	研究疾病	使用洗脱期时长（年）	直接限定法	一致性比较法	逆向生存函数法	具体指标
Sascha Abbas	2012	糖尿病、结直肠癌、心衰	1-8		√		发病高估率
J. D. Stein	2013	青光眼、黄斑变性	3-5		√		发病高估率
M. Rosenlund	2020	24 中慢性病	0.5-9		√		发病高估率
N. Schmedt	2017	多发性硬化	8		√		发病高估率
R. Camplain	2017	心衰	0.5-3		√		发病高估率
K. M. Ødegaard,	2022	心衰	1-8		√		发病高估率
Alain Vanasse	2012	精神分裂	7		√		Kappa、PPVs

E. I. Benchimol,	2014	慢性肠道疾病	4		√		灵敏度、准确度、PPVs、阴性预测值
N. Beaudet	2014	下背痛	7		√	√	逆向生存函数、PPVs、Kappa
S. Asghari	2009	糖尿病	5		√	√	逆向生存函数、Kappa
K. J. Brameld	2003	糖尿病、心肌梗死	13			√	逆向生存函数
L.C.Lin	2014	糖尿病视网膜病变	3	√			文献检索
C. Ohlmeier	2019	亨廷顿氏病	4	√			文献检索
B. K. Koo	2014	糖尿病、动脉粥样硬化	2	√			文献检索
M. Jürisson	2014	髋骨骨折	1	√			直接给定
A. S. Bukhbinder	2022	阿尔兹海默	6	√			直接给定
L. Zhou	2022	血栓形成	1	√			直接给定
J.Katzenellenbogen	2019	风湿性心脏病	13.5	√			数据限制
Ou, Y	2012	青光眼、阿尔兹海默症、其他痴呆	3	√			文献检索
B. B. Clissold	2017	中风	5	√			文献检索
J. K. Ahn	2020	风湿性疾病	3	√			直接给定
K. K. Anderson	2012	精神分裂	4-6	√			直接给定
J.Katzenellenbogen	2010	心肌梗死	15	√			数据限制
Gerhard Sulo	2015	心肌梗死	3,5,7,8,10	√			直接给定
J. Czwikla	2017	乳腺癌、前列腺癌、结直肠癌	1—7	√			数据限制
Rassen, J. A.	2019	慢阻肺、银屑病、精神分裂、囊性纤维症	1—4	√			数据限制

3. 讨论

本次研究系统梳理了使用行政管理数据库确定慢性疾病发病病例，限定最佳洗脱期时长的方法。目前，半数左右的研究（n=26，49.15 %）在确定慢性疾病发病病例时，使用了洗脱期，但洗脱期的时长目前尚没有统一结论。纳入研究的 26 篇文献，涉及了 24 种疾病，采用的洗脱期时长从 0.5-15 年不等。确定洗脱期时长的方法主要有三种：直接限定法、一致性比较法和逆向生存曲线法，其中最常用的还是直接限定法，逆向生存曲线法用的相对较少。

本次研究发现半数（n=13，50.00%）的研究使用直接限定法给定洗脱期，该方法利用专家建议、文献检索等方式直接限定洗脱期长短。此方法虽便捷，易获得，但缺乏量化的数据支持，一定程度上缺乏说服力，研究者也在文章局限性中指出，直接限定的洗脱期可能会导致发病估计不准确[11, 12]；2 篇文献选择直接使用最长数据获得年份为洗脱期，导致数据无法有效利用，通常只能评估一年的发病情，研究信息获取不足。部分文献（n=9，34.62%）使用一致性检验的相关指标确定最佳洗脱期。不同的文献使用的一致性检验指标不同，目前使用较多的一般为 Kappa 值和高估率，其中 Kappa 值的判定标准采用 Byrt、Shabnam[7]等人 1993 年提出的分类标准进行判断，其余一致性指标没有具体的判断依据。多数研究选择 Kappa 值大于 0.8（分类为“一致性好”）作为选择标准[9, 10, 13]，但有研究认为 在确定高患病率的疾病（如糖尿病、心血管疾病）发病病例时，高的 Kappa 值可能与病例的高患病人数有关，稳定性有待进一步探究[10]。其他一致性指标则没有具体的定量限定标准，如 Sascha Abbas[6]等人采用 10%的高估率为临界标准，Mats Rosenlund[4]等人采用 20%；对于 PPVs，有研究选择 80%为评判指标[13]，而 Eric I. Benchimol 等，则选择准确度达到 90%作为评判标准[14]。仅有少部分研究（n=3，11.54%）借用生存函数的思路，使用逆向生存函数确定最佳洗脱期。在 Kate J Brameld[8]等人于 2003 首次提出逆向生存函数法确定洗脱期后，该方法被陆续使用[9, 10]。逆向生存函数法借用生存函数的思路，绘制生存函数和风险函数曲线，并依据生存概率和风险概率对最佳洗脱期时长进行确定。研究者认为该方法可以最大限度利用数据，使用所有能获得的记录估计生存函数。同时，通过风险函数和风险函数曲线可以直观发现生存概率和风险值的集中趋势，有效定量判断最佳的洗脱期时长。但该方法也存在一定的局限性，即判断标准尚未完全统

一。如 Kate J Brameld[5] 等人认为当风险小于 0.00001 时，即达到最佳洗脱期时长，按照此标准下，糖尿病的洗脱期长达 13 年。Asghari[10]等人并未使用这一标准，而是仅通过生存概率不再变化这一指标，确认糖尿病的洗脱期为 5 年，使得同一疾病洗脱期时长有较大差异。值得注意的是，由于洗脱期时长确定方法尚未统一，且判断标准也没有达成共识，因此许多研究者在探究最佳洗脱期时，选择计算多个一致性指标，结合逆行生存函数，进行综合判断[8-10]。

在洗脱期时长方面，研究纳入的文献采用的洗脱期从 0.5-15 年不等，不同疾病使用的洗脱期时长差异较大。有研究认为，该现象是由于不同疾病有不同的发展轨迹和疾病特征导致[6]。本次研究同时发现，即使同一疾病，如糖尿病，洗脱期时长差异也较大。我们认为，造成该差异的原因，可能是由于数据的可使用年限、人群就医习惯、数据收集方式等差异导致。因此，我们认为，在使用不同数据库确认不同疾病的发病病例时，均应选择合适的方法探索最佳洗脱期，并在此基础上进行进一步疾病流行病学特征的研究，这一观点也与一些研究者不谋而合[2, 3, 15]。这提示我们，在后续研究者使用中国医保、疾病注册等行政管理数据进行慢性疾病的流行病特征相关分析时，考虑医保政策、环境人种等差异存在，不建议直接应用其他国家的洗脱期，应在数据支持的情况下，选择合适的方法确定最佳洗脱期时长，并在此基础上进行后续分析研究。

行政管理数据（如医保数据、注册数据等）作为宝贵的数据资源，涵盖范围广，年份跨度大，获取便捷，相对于传统的调查有一定的优势。中国自 2012 年改革医保制度[16, 17]以来，大力推广医保政策，扩大涵盖人群范围，至今已超 10 年，有效数据积累年份逐渐丰富。目前在万方、知网、维普检索的利用医保、医院注册数据等行政管理数据进行流行病相关分析的文章仅 21 篇，远少于行政管理数据完备，研究范式成熟的其他国家或地区，说明中国在此方面的应用仍处于探索阶段，未来发展潜力巨大。同时，中国在使用医保数据库进行分析时，洗脱期的概念尚未得到重视，检索到的 21 篇文献均未使用洗脱期，同样，也未有文献聚焦于利用中国人群医疗保险数据进行疾病流行病学特征分析时洗脱期时长的研究。因此，本此研究系统总结了三种洗脱期确定的方法，为后续研究者利用中国医保数据，准确识别中国人群慢性疾病新发病例，探究疾病流行病学特征提供相关思路和方法支持。

作者贡献

杨文怡提出研究选题方向，负责文献检索收集和整理，并撰写论文初稿；王敬鑫进行文献数据提取和提取信息整理；艾丽梅负责论文的修订；万霞负责文章的质量控制及审校，对文章整体负责；所有作者确认了论文的最终稿。

利益冲突情况

本文无利益冲突。

参考文献

[1] MENG Q, XU L, ZHANG Y, et al. Trends in access to health services and financial protection in China between 2003 and 2011: a cross-sectional study [J]. Lancet, 2012, 379(9818): 805-14. DOI: 10.1016/s0140-6736(12)60278-5

[2] ØDEGAARD K M, LIRHUS S S, MELBERG H O, et al. A nationwide registry study on heart failure in Norway from 2008 to 2018: variations in lookback period affect incidence estimates [J]. BMC Cardiovasc Disord, 2022, 22(1): 88. DOI: 10.1186/s12872-022-02522-y

[3] CZWIKLA J, JOBSKI K, SCHINK T. The impact of the lookback period and definition of confirmatory events on the identification of incident cancer cases in administrative data [J]. BMC Med Res Methodol, 2017, 17(1): 122. DOI: 10.1186/s12874-017-0407-4

[4] ROSENLUND M, EKSTROM N, TORNBLOM M, et al. Impact of variable look-back periods on the incidence rates of chronic diseases using real world data [J]. Pharmacoepidemiology and Drug Safety, 2020, 29(9): 1086-92. DOI: 10.1002/pds.5066

[5] SCHMEDT N, KHIL L, BERGER K, et al. Incidence of Multiple Sclerosis in Germany: A Cohort Study Applying Different Case Definitions Based on Claims Data [J]. Neuroepidemiology, 2017, 49(3-4): 91-8. DOI: 10.1159/000481990.

[6] ABBAS S, IHLE P, KOESTER I, et al. Estimation of Disease Incidence in Claims Data Dependent on the Length of Follow-Up: A Methodological Approach [J]. Health Services Research, 2012, 47(2): 746-55. DOI: 10.1111/j.1475-6773.2011.01325.x.
[7] BYRT T, BISHOP J FAU - CARLIN J B, CARLIN J B. Bias, prevalence and kappa [J]. J Clin Epidemiol. 1993;46(5):423–429. doi: 10.1016/0895-4356(93)90018-V.
[8] BRAMELD K J, HOLMAN C D, LAWRENCE D M, et al. Improved methods for estimating incidence from linked hospital morbidity data [J]. Int J Epidemiol, 2003, 32(4): 617-24. DOI: 10.1093/ije/dyg191.
[9] BEAUDET N, COURTEAU J, SARRET P, et al. Improving the selection of true incident cases of low back pain by screening retrospective administrative data [J]. Eur J Pain, 2014, 18(7): 923-31. DOI: 10.1002/j.1532-2149.2013.00437.x
[10] ASGHARI S, COURTEAU J, CARPENTIER A C, et al. Optimal strategy to identify incidence of diagnostic of diabetes using administrative data [J]. BMC Med Res Methodol, 2009, 9(62. DOI: 10.1186/1471-2288-9-62
[11] AHN J K, HWANG J, SEO G H. Incidence and risk of developing rheumatic diseases in 19,724 patients with palindromic rheumatism in South Korea: A nationwide population-based study [J]. Joint Bone Spine, 2021, 88(3): 105128. DOI: 10.1016/j.jbspin.2020.105128.
[12] OHLMEIER C, SAUM K-U, GALETZKA W, et al. Epidemiology and health care utilization of patients suffering from Huntington's disease in Germany: real world evidence based on German claims data [J]. Bmc Neurology, 2019, 19(1): 318. DOI: 10.1186/s12883-019-1556-3.
[13] VANASSE A, COURTEAU J, FLEURY M J, et al. Treatment prevalence and incidence of schizophrenia in Quebec using a population health services perspective: different algorithms, different estimates [J]. Soc Psychiatry Psychiatr Epidemiol, 2012, 47(4): 533-43. DOI: 10.1007/s00127-011-0371-y.
[14] BENCHIMOL E I, GUTTMANN A, MACK D R, et al. Validation of international algorithms to identify adults with inflammatory bowel disease in health administrative data from Ontario, Canada [J]. J Clin Epidemiol, 2014, 67(8): 887-96. DOI: 10.1016/j.jclinepi.2014.02.019.
[15] STEIN J D, BLACHLEY T S, MUSCH D C. Identification of persons with incident ocular diseases using health care claims databases [J]. Am J Ophthalmol, 2013, 156(6): 1169-75.e3. DOI: 10.1016/j.ajo.2013.06.035.
[16] 谢莉琴， 秦盼盼， 高星， et al. 中国城乡居民基本医疗保险制度发展历程、挑战与应对策略 [J]. 中国公共卫生, 2020, 36(12): 1673-6.
XIE L Q, QIN P P, GAO X, et al.Development, challenges and strategies of basic medical insurance system for urban and rural residents in China [J]. Chinese Journal of Public Health. 2020, 36(12): 1673-6. DOI:10.11847/zgggws1126456
[17] WU R, LI N, ERCIA A. The Effects of Private Health Insurance on Universal Health Coverage Objectives in China: A Systematic Literature Review [J]. Int J Environ Res Public Health, 2020, 17(6): 1660-4601. DOI: 10.3390/ijerph17062049